Metacognition and Learning
https://doi.org/10.1007/s11409-020-09225-9

Confident or familiar? The role of familiarity ratings in adults' confidence judgments when estimating fraction magnitudes

Charles J. Fitzsimmons[1] · Clarissa A. Thompson[1] · Pooja G. Sidney[2]

Received: 14 June 2019 / Accepted: 23 April 2020/Published online: 15 May 2020

Abstract
Understanding fraction magnitudes is especially important in daily life, but fraction reasoning is quite difficult. To accurately reason about fraction magnitudes, adults need to monitor what they know and what they do not know. However, little is known about which cues adults use to monitor fraction performance. Across two studies, we examined adults' trial-by-trial fraction estimates, confidence judgments, and ratings of fraction familiarity. Adults were more confident when their estimates were more precise as well as when estimating fractions they rated as more familiar. However, adults judged their confidence in estimating fraction magnitudes, in part, based on their familiarity with each fraction. The role familiarity cues play in judgments of confidence with fractions suggests that people may be less likely to check for errors when reasoning about highly-familiar fractions.

Keywords Fractions · Metacognition · Confidence judgments · Whole number bias · Familiarity · Cue utilization

Across multiple countries, students' understanding of fraction magnitudes is predictive of their overall mathematics performance, even when controlling for socio-economic background, working memory, general intelligence, and fraction arithmetic abilities (Siegler et al. 2012; Torbeyns et al. 2015). However, compared to whole numbers, children and adults are less

Part of this report was accepted as a conference proceeding to the North American Chapter of the International Group for the Psychology of Mathematics Education (PME-NA, 2019). The authors retain full copyrights of the submission. Pre-registrations and project information can be viewed here: https://osf.io/4uygd/.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11409-020-09225-9) contains supplementary material, which is available to authorized users.

✉ Charles J. Fitzsimmons
cfitzsi4@kent.edu

[1] Psychological Sciences, Kent State University, Kent, OH 44240, USA

[2] Department of Psychology, University of Kentucky, Lexington, KY 40506, USA

precise when estimating the magnitude of fractions (e.g., Siegler and Opfer 2003; Siegler et al. 2011). This is particularly troublesome in the case of adults because they report that fractions are important for their work (Handel 2016), and rational number understanding is critical to many health and financial decisions that adults make on a day-to-day basis (e.g., Peters et al. 2007; Peters et al. 2006). One source of difficulty in understanding fraction magnitudes is the whole number bias: when knowledge of whole number concepts interferes with fraction reasoning (Alibali and Sidney 2015; Ni and Zhou 2005).

Whole number bias results from focusing on the whole number numerator or denominator components of fractions *independently*, as compared to the ratio of the two components, and overextending whole number reasoning when not appropriate for fraction tasks. For example, children tend to estimate fractions with larger components (e.g., 15/30) as larger than equivalent fractions with smaller components (e.g., 1/2), presumably because 15 and 30 are larger than 1 and 2, even though 15/30 = 1/2 (Braithwaite and Siegler 2017). One account of whole number bias argues that these errors arise because children receive formal instruction and experience with whole numbers that emphasizes the differences, rather than similarities, between these two types of numbers (e.g., Siegler et al. 2011; Ni and Zhou 2005). These errors often persist into adulthood (e.g., Alibali and Sidney 2015; Ni and Zhou 2005; Opfer and DeVries 2008; Sidney et al. 2018), but it is an open question whether adults are *aware* of whole number bias when reasoning about fraction magnitudes. In other words, are adults able to metacognitively monitor their fraction performance?

Can adults accurately monitor their fraction estimation performance?

Awareness of performance via self-monitoring is important for academic achievement because monitoring influences control of study behaviors (e.g., Dunlosky and Rawson 2012; Metcalfe 2009). For example, individuals may be more likely to notice and correct errors when reasoning about fraction magnitudes if they can monitor their understanding that fractions with larger components do not always have larger magnitudes than fractions with smaller components (e.g., 15/30 is not >1/2 even though 15 > 1 and 30 > 2). As this example illustrates, monitoring may be especially important when reasoning about difficult fraction concepts, a domain in which individuals harbor known misconceptions (Alibali and Sidney 2015; Ni and Zhou 2005), often automatically misapply familiar whole number concepts (Siegler et al. 2011), and may need to inhibit this prior knowledge (Siegler and Pyke 2013).

Recent work has investigated how people monitor their performance in numerical tasks (Nelson and Fyfe 2019; O'Leary and Sloutsky 2016, 2018; Wall et al. 2016) by using their confidence to discriminate among items or tasks on which they perform better or worse (e.g., Bjork et al. 2013; Dunlosky and Metcalfe 2008). That is, people's relative confidence from one item to the next may or may not align well with their relative performance from one item to the next. Thus, assessing people's confidence when reasoning about fractions provides insights into how they monitor their understanding.

Cues adults use when making fraction confidence judgments

Ideally, only an individual's ability to generate an accurate answer would affect their confidence in their responses. However, according to the cue-utilization approach (Koriat 1997),

people use a variety of different heuristics and cues to make metacognitive judgments (e.g., Alter and Oppenheimer 2009; Hertzog et al. 2003; Koriat 2008; Koriat and Levy-Sadot 2001; Reder and Ritter 1992). People's judgments will be accurate to the extent that the cues utilized are predictive of their performance. Some cues, such as processing fluency or familiarity, may be predictive of performance in some cases, yet lead to less accurate judgments in others. However, little work has evaluated whether individuals use their familiarity with numbers as a cue to make concurrent confidence judgments of their math performance.

Familiarity ratings may be impacted by environmental input

Individuals may use their familiarity with numbers to make confidence judgments when they estimate whole number magnitudes. For example, in one study of children's whole number magnitude estimation (Wall et al. 2016), *most* children were more confident and precise in their estimates within smaller, more familiar numerical ranges (e.g., 0–100) than larger, unfamiliar numerical ranges (e.g., 0–1000). However, *some* children's confidence was less well-aligned: they were more confident in their small-range estimates even when they were *equally* precise in smaller and larger ranges. Others were more confident in the smaller range even when their estimates in this range were *less* precise. Overall, students may feel more confident with smaller numbers due to familiarity with the 0–100 range because smaller numbers are encountered more frequently within the environment than larger numbers (Dehaene and Mehler 1992). Similar to the variability in environmental input of whole numbers, evidence suggests that some fractions are also encountered more frequently than others in common math textbooks (Braithwaite and Siegler 2018). When the common fractions from textbooks are provided as input to computational models, the models exhibit errors similar to those of children (Braithwaite et al. 2019; Braithwaite et al. 2017). It is an open question as to whether the frequency of encountering fractions in textbooks has downstream consequences for adults' familiarity with fractions and whether this familiarity serves as a cue that influences adults' fraction confidence judgments.

Current studies

In the current studies, we examined adults' confidence, familiarity, and number line estimation precision (i.e., how close their estimate was to the true location of the fraction) for equivalent fractions with smaller (e.g., 1/2) and larger components (e.g., 15/30). We manipulated the size of components because fractions with larger components occur less frequently in math textbooks (Braithwaite and Siegler 2018), components affect how people process fraction magnitudes (i.e., whole number bias; Ni and Zhou 2005), and little is known about the impact that components have on familiarity and confidence. By examining the role of familiarity in adults' confidence judgments, we provide an important first step in evaluating the information that people use to monitor their ability to reason about fractions. Specifically, examining familiarity allows us to evaluate how one heuristic may inform adults' math metacognition. Adults who use familiarity to inform their confidence during fraction reasoning may be less likely to check for errors during financial or health decisions. For example, Peters et al. (2019) found that adults with high subjective confidence in their numerical abilities (e.g., "How good are you at working with fractions?") and low objective numerical abilities (i.e., scores on a

measure involving items that might be found on a math test) were more likely to have lower self-reported financial outcomes and higher Lupus disease activity compared to individuals whose confidence aligned with their objective numeracy scores. However, this research did not examine adults' trial-by-trial confidence nor did it examine factors that might contribute towards adults' confidence in their performance.

We hypothesized that adults would be more precise (H1a), confident (H1b), and familiar (H1c) with fractions that have smaller components compared to those with larger components (https://osf.io/4uygd/). Of primary interest, we asked whether confidence judgments made immediately after fraction number line estimates were more strongly related to the precision of the estimates, familiarity with the particular fractions, or equally related to both. We hypothesized that confidence judgments would be related to both estimation precision as well as familiarity (H2). However, given the dearth of research examining familiarity with fractions, we did not have a specific hypothesis about which relation would be stronger. Relying on one's familiarity with fractions instead of performance to make confidence judgments might result in misallocating future study time and failing to detect errors when reasoning about fractions.

Method

Participants

We recruited two community-based samples from Amazon's Mechanical Turk as part of a larger study investigating adults' fraction reasoning (Fitzsimmons et al. in press). We recruited each sample at different time points, and participants from the first experiment were excluded from participating in the second. Although there are some methodological differences between the two experiments, both include the focal tasks for this report (i.e., number line estimation, confidence judgments, and familiarity ratings). Thus, we report both experiments together, and Experiment 2 serves as a replication of Experiment 1. In both experiments, we aimed to recruit 110 participants based on a power analysis to detect a small effect with 80% power and 10% attrition. We powered both experiments to replicate the smallest effect reported by Braithwaite and Siegler (2017) because this was the primary aim of the larger study on fraction reasoning.

We screened participants' responses to ensure they followed directions for each task using pre-registered inclusion criteria (https://osf.io/4uygd/). For instance, we excluded participants from analyses if they self-reported calculator use and only made estimates on 10% of the number line. Of the adults who met pre-registered inclusion criteria (Experiment 1: M_{age} = 35.06 years, SD = 10.25 years, age range: 21–69 years; Experiment 2: M_{age} = 34.96, SD = 11.35 years, age range: 19–71 years), most were white and nearly half were college graduates. See Table 1 for demographics from each sample.

Tasks and procedure

Participants completed all tasks within an online Qualtrics survey at their own pace on their personal electronic device. In both experiments, participants estimated the location of a fraction on a number line then provided a confidence judgment immediately after each trial. Participants rated their confidence immediately after each trial to increase the likelihood that

Table 1 Demographics from Experiments 1 (E1) and 2 (E2)

Demographics	Frequency (percent)	
Gender	E1	E2
Female	34 (37.4)	42 (41.6)
Male	56 (61.5)	59 (58.4)
Did not report	1 (1.1)	–
Total	91	101
Race		
White	66 (72.5)	65 (64.4)
Black or African American	9 (9.9)	10 (9.9)
American Indian or Alaskan Native	4 (4.4)	2 (2.0)
Asian	4 (3.3)	13 (12.9)
Hispanic or Latino	1 (1.1)	5 (5.0)
Multiple races	6 (6.6)	4 (4.0)
Did not report	1 (1.1)	2 (2.0)
Education		
Completed high school	17 (18.7)	12 (11.9)
Post high school other than college	1 (1.1)	3 (3.0)
Some college	22 (24.2)	25 (24.8)
College graduate	41 (45.1)	47 (46.5)
Postgraduate	9 (9.9)	11 (10.9)
Did not report	1 (1.1)	3 (3.0)

they would base their confidence judgments on their estimation performance–a particularly diagnostic cue. At the end of both experiments, participants provided familiarity ratings for each fraction presented during the number line estimation task. That is, familiarity was the *last* task participants completed.

There are three primary differences between experiments stemming from the research questions of the parent study (Fitzsimmons et al. in press). First, in Experiment 1 we included measures of numerical inhibition and updating. Second, the between-participants' manipulation differed between studies: in Experiment 1, half of the participants estimated under a time constraint and the other half did not; in Experiment 2, number line estimation was completed either before or after a fraction equivalence task. Finally, in Experiment 2 we measured participants' self-reported strategy use. Because participants reported their strategy use after each number line trial, we reduced the number of estimation trials to ensure that the length of the overall experimental session was manageable.

We control for condition (the presence of the time constraint in Experiment 1 and order of equivalence task in Experiment 2) in our analyses for both experiments, but these experimental manipulations were not central to the results of the current study. For example, in Experiment 1 confidence was lower in the timed compared to non-timed condition, but there was no difference in the intra-individual relation between confidence and estimation precision or between confidence and familiarity in either condition. As we are interested in whether estimation precision or familiarity is more strongly related to confidence, we do not report conditional differences in estimation performance.

Number line estimation Participants estimated the location of 44 fractions in Experiment 1 and 12 fractions in Experiment 2, one at a time, on number lines with 0 located at the left endpoint and 1 located at the right (see the Appendix Table 5 for all fraction stimuli). As noted

above, we reduced the number of estimation items in Experiment 2 because participants reported their strategies for each trial which is a time-intensive process. The 44 fractions in Experiment 1 consisted of 11 magnitudes in the 0–1 range. Each magnitude was represented by four different equivalent fractions that had smaller (e.g., 1/2 or 2/4) or larger (e.g., 12/24 or 15/30) components. The 12 fractions in Experiment 2 consisted of six magnitudes that spanned the 0–1 range. Each magnitude was represented by a fraction with smaller (e.g., 1/6) or larger (e.g., 12/72) components. Both sets of fractions were adapted from prior work (Braithwaite and Siegler 2018) and designed to examine the effects of whole number components on fraction reasoning. Braithwaite and Siegler (2017) created smaller component fractions which were in their simplified form (e.g., 1/2) or that had been multiplied by a number, such as 2/2, to create a non-simplified, smaller-component fraction. Larger-component fractions were created by multiplying the simplified fraction by a larger number, such as 6/6. Thus, component size is operationalized relative to the simplified form of the fraction.

For each trial, adults were asked, “Where does this fraction go on this number line (0 to 1)?” with the to-be-estimated fraction located above the line. Adults could click on the line or slide an icon to indicate their estimate. To evaluate number line estimation precision, we calculated the absolute difference between the provided estimate and actual fraction magnitude and multiplied these values by 100 to transform them into percent absolute error (PAE; see Siegler and Booth 2004). PAE is calculated for each individual participant, for each trial, and then averaged across all trials. PAE is inversely related to estimation precision such that higher values of PAE indicate less precise estimates.

Confidence judgments Immediately after each number line estimation trial, participants rated whether they were “not so sure,” “kind of sure,” “pretty sure,” or “totally sure” that they placed the fraction at *approximately* the correct location on the line (Wall et al. 2015). The participants’ number line estimates were not visible on the screen when they made their confidence judgment, but they were reminded of the fraction that they had just estimated.

We chose to use this 4-point scale because it had been previously used in research with adults who rated their confidence in their number line estimation performance. Thus, we can compare our results to the existing literature on adults’ confidence judgments (Wall et al. 2015). We chose to use a 4-point rather than a 3-point scale to reduce the likelihood that participants would provide invariant confidence ratings. When there is no variability in a participant’s confidence ratings, we are unable to calculate an intra-individual correlation, an issue observed in the Wall et al. (2016) experiment. Furthermore, research in metamemory suggests that 4-point, 20-point, and 100-point scales lead to similar levels of metacognitive precision (e.g., Tekin et al. 2018; Tekin and Roediger, 2017). For example, Tekin and Roediger (2017) found that the relation between confidence and accuracy on a word-identification task or a face identification task was comparable between 4-, 5-, 20-, and 100-point scales. That is, confidence ratings on different scales led to similar levels of confidence-judgment accuracy. Similar results were found when comparing verbal to numeric scales (Tekin et al. 2018). We acknowledge that measuring estimation precision and confidence on different scales limits our ability to make claims about over- and under-confidence (i.e., calibration). However, we were primarily interested in which cues adults used to discriminate between items that they performed more or less well on (i.e., relative confidence).

Familiarity judgments Participants rated their fraction familiarity on a six-point scale from “not familiar at all” to “very familiar” for each of the fractions that they had previously

estimated in the number line task similar to a word familiarity rating scale used in past work (Dumas et al. 2002). Only the end-points of the scale were labeled. Participants were asked to "Please rate how familiar you are with each fraction on the list, i.e. the extent to which you have seen the fraction before and feel comfortable using the fraction." All fractions were visible together on the screen as participants made their familiarity judgments at the end of the experiment after all other tasks were completed to limit any potential influence judging familiarity might have on making confidence judgments.

Results

Number line precision, confidence, and familiarity

We evaluated estimation precision by calculating participants' average percentage of absolute error (PAE) as well as the precision of smaller and larger component fractions, separately. See Table 2 for group-level means, standard deviations, and correlations averaged across smaller- and larger-component fractions for both experiments.

We predicted that adults would be more precise (H1a) when estimating fractions with smaller compared to larger components. In Experiment 1 this hypothesis was supported: adults were more precise (i.e., lower PAE) when estimating fractions with smaller ($M = 8.15\%$, $SD = 8.14\%$) compared to larger components ($M = 9.26\%$, $SD = 8.05\%$), $F(1, 89) = 10.79$, $p = .001$, $\eta_p^2 = .11$. In contrast, adults were equally precise when estimating fractions with smaller ($M = 7.37\%$, $SD = 8.45\%$) and larger ($M = 7.74\%$, $SD = 8.22\%$) components in Experiment 2, $F(1, 99) = 0.56$, $p = .456$, $\eta_p^2 = .01$ (see Table 3 for descriptive statistics from each study).

To examine the robustness of the effect of component size on estimation precision, we conducted an analysis of the combined data from both experiments in a linear-mixed effects model using the *lme4* package (Bates et al. 2015) as implemented in R (R Core Team 2015). More information about this model can be found in the Supplemental File. There was a significant effect of component size, $F(1, 5024) = 8.72$, $p = .003$. When averaged across experiments, adults were more precise when estimating fractions with smaller compared to larger components, $b = 0.82$, $t(5024) = 2.95$, $p = .003$.

We also predicted that adults would be more confident (H1b) when estimating fractions that had smaller compared to larger components. As expected, adults were more confident when estimating fractions with smaller ($M = 2.92$, $SD = 0.60$) compared to larger ($M = 2.75$, $SD = .66$) components in Experiment 1, $F(1, 89) = 59.94$, $p < .001$, $\eta_p^2 = .40$. This effect replicated

Table 2 Group level means, standard deviations, and correlations for Experiments 1 and 2

		PAE	Confidence	Familiarity
	M (SD)	7.55% (7.98%)	2.76 (.64)	4.04 (1.10)
PAE	8.71% (7.93%)	–	−.46**	−.21*
Confidence	2.83 (.62)	−.52**	–	.43**
Familiarity	4.10 (1.07)	−.37**	.52**	–

Note. PAE = percent absolute error averaged across smaller- and larger-component fractions. Confidence was rated on a four-point scale, and familiarity was rated on a six-point scale. Experiment 1 correlations are below the diagonal, and Experiment 2 correlations are above the diagonal. * $p < .05$ ** $p < .01$

Table 3 Descriptive statistics for PAE, confidence, and familiarity in Experiment 1 and 2

Experiment			Smaller Component (SD)	Larger Component (SD)
1	H1a	Percent Absolute Error	8.15% (8.14%)	9.26% (8.05%)
	H1b	Confidence Judgments	2.92 (.60)	2.75 (.66)
	H1c	Familiarity Ratings	4.45 (.90)	3.75 (1.30)
2	H1a	Percent Absolute Error	7.37% (8.45%)	7.74% (8.22%)
	H1b	Confidence Judgments	2.85 (.67)	2.68 (.66)
	H1c	Familiarity Ratings	4.75 (1.08)	3.34 (1.46)

in Experiment 2: adults were more confident when estimating fractions with smaller ($M = 2.85$, $SD = 0.67$) compared to larger ($M = 2.68$, $SD = 0.66$) components, $F(1, 99) = 22.69$, $p < .001$, $\eta_p^2 = .19$.

Similarly, as predicted (H1c), we found that adults were more familiar with fractions that had smaller compared to larger components in both experiments. In Experiment 1, adults were more familiar with fractions that had smaller ($M = 4.45$, $SD = 0.90$) compared to larger ($M = 3.75$, $SD = 1.30$) components, $F(1, 89) = 113.32$, $p < .001$, $\eta_p^2 = .56$. This replicated in Experiment 2: adults were more familiar with fractions that had smaller ($M = 4.75$, $SD = 1.08$) compared to larger ($M = 3.34$, $SD = 1.46$) components, $F(1, 99) = 108.51$, $p < .001$, $\eta_p^2 = .52$. Thus, the size of the whole number components impacted adults' estimation precision, confidence, and familiarity with equivalent fractions. Even when adults' estimates of equivalent fractions were equally precise in Experiment 2, they were still more confident and familiar with fractions that had smaller components. Adults' higher confidence when estimating more familiar fractions could be due to additional experience with some fractions compared to others.

Linking familiarity to experience

If experience with fractions drives the differences in confidence and familiarity between smaller- and larger-component fractions, then fractions commonly found in textbooks should be rated as more familiar due to these frequent encounters. Thus, we examined whether the frequently encountered fractions presented in Experiment 1 (1/4, 1/3, 1/2, 3/4, 2/3; a subset of common fractions identified by Braithwaite and Siegler 2018) were indeed rated as more familiar than other simplified, small-component fractions not identified as common across different textbooks. In three, between-within mixed ANOVAs controlling for condition, we compared adults' familiarity, confidence, and PAE for these frequently encountered fractions to the other simplified fractions (1/5, 2/9, 3/7, 5/9, 4/5, 5/6). As expected, adults rated these common fractions as more familiar ($M = 5.43$, $SD = .60$) than the other simplified fractions ($M = 4.03$, $SD = 1.17$), $F(1, 89) = 199.94$, $p < .001$, $\eta_p^2 = .69$. Furthermore, adults felt more confident in their estimates of common fractions ($M = 3.23$, $SD = .62$) than other simplified fractions ($M = 2.77$, $SD = .67$), $F(1, 89) = 90.75$, $p < .001$, $\eta_p^2 = .51$. Despite being more confident with these more familiar fractions, adults were no more precise when estimating common ($M = 8.0\%$, $SD = 10.43\%$) than less common ($M = 8.0\%$, $SD = 9.0\%$) simplified fractions, $F(1, 89) = .003$, $p = .96$. Thus, environmental experiences can impact subjective familiarity and confidence in ways that may be misleading for adults' fraction metacognition. In both experiments, adults were more confident when they estimated fractions that they also rated as more familiar despite being equally precise in their estimates. In Experiment 1, this was the case when comparing frequently encountered small-component fractions to less

frequently encountered small-component fractions. In Experiment 2, this was the case when comparing equivalent fractions with smaller or larger components.

Which cues relate to adults' confidence judgments?

To examine whether adults' estimation precision or familiarity with fractions was more strongly related to their judgments of confidence, we calculated gamma correlations for each participant. Gamma is a commonly used measure for assessing intra-individual variability in trial-to-trial monitoring abilities, and it accounts for the ordinal nature of confidence judgments (Nelson 1984; Wall et al. 2016).[1] Gamma is a measure of *relative* confidence. That is, are items that participants rate with higher confidence related to better performance or higher familiarity relative to items in which participants have lower confidence? Furthermore, because gamma is an interval measure (range − 1 to 1) and is normally distributed in our dataset (see Table 4), these values meet the assumptions of parametric tests and can be analyzed using measures such as ANOVAs and t-tests (see Nelson 1984 for more information).

First, we evaluated whether confidence was related to estimation precision and familiarity by calculating the intra-individual gamma correlation between adults' confidence and estimation precision (PAE) and between confidence and familiarity. We chose to calculate relative confidence because we were interested in whether adults' estimation precision or fraction familiarity was more strongly related to their confidence judgments. We report inverse values of gamma for the relation between confidence and PAE for ease of interpretation because lower values of PAE indicate more precise estimates. That is, we multiplied the gamma values for the CJ-PAE relation by −1 so that higher values indicate higher confidence is related to more precise estimates.

We predicted that confidence judgments would relate to estimation precision and familiarity (H2). This hypothesis was supported in both experiments. In Experiment 1, gamma for the CJ-PAE relation was positive ($M = 0.34$, $SD = 0.25$) and significantly different from zero, $t(90) = 12.88$, $p < .001$, 95% CI [0.28, 0.39], $d = 1.35$. The gamma for the CJ-familiarity relation was also positive ($M = 0.51$, $SD = 0.38$) and significantly different from zero $t(88) = 12.64$, $p < .001$, 95% CI [0.43, 0.59], $d = 1.34$. This replicated in Experiment 2. The gamma for the CJ-PAE relation was positive ($M = 0.14$, $SD = 0.49$) and significantly different from zero, $t(84) = 2.57$, $p = 0.012$, 95% CI [0.03, 0.24], $d = 0.28$, as was the gamma for the CJ-familiarity relation ($M = 0.31$, $SD = 0.58$), $t(77) = 4.78$, $p < .001$, 95% CI [0.18, 0.45], $d = 0.54$. Thus, in both experiments adults were more confident when their estimates were more precise as well as when they estimated fractions they rated as more familiar. Therefore, *both* estimation precision and subjective familiarity relate to adults' judgments of confidence when estimating fraction magnitudes.

Second, and of central interest, we compared adults' intra-individual gamma for the CJ-PAE relation to the CJ-familiarity relation in an ANOVA for each experiment. As with our prior analyses, we control for the between-subjects experimental manipulations in each study. We found that the relation between confidence judgments and familiarity was stronger than the relation between confidence and PAE in both Experiment 1, $F(1, 87) = 18.07$, $p < .001$, 95%

[1] We acknowledge that there is an ongoing debate as to which measure of metacognitive accuracy is best (Fleming and Lau 2014; Higham and Higham 2019). However, we chose to calculate gamma given the similarities between gamma and other signal-detection measures of metacognitive sensitivity (Higham and Higham 2019). Furthermore, gamma is appropriate for our continuous measure of number line estimation precision, whereas signal-detection analyses such as AROC require a dichotomized response (e.g., correct/incorrect), and PAE is a continuous measure of performance.

Table 4 Descriptive statistics including normality information for gammas in each experiment

Experiment		Gamma (SD)	Skew	Kurtosis
1	CJ to PAE	0.34 (.25)	−0.71	2.14
	CJ to familiarity	0.51 (.38)	−1.78	4.48
2	CJ to PAE	0.14 (.49)	−0.31	−0.19
	CJ to familiarity	0.31 (.58)	−0.66	−0.37

CI for the difference in gammas [.09, .24], $\eta_p^2 = 0.17$, and Experiment 2, $F(1, 76) = 7.08$, $p = .009$, 95% CI for the difference in gammas [.05, .38], $\eta_p^2 = 0.09$ (see Fig. 1). Thus, in both experiments, participants' familiarity was more strongly related to their confidence than their PAE, a measure of their objective estimation performance.

Discussion

In two experiments, we examined adults' ability to monitor their number line estimation performance as they estimated equivalent fractions with smaller and larger components. In line with our hypotheses, adults from both experiments were less confident and less familiar with larger-component fractions relative to smaller-component fractions. In Experiment 1, adults' confidence judgments reflected their performance: they were less precise when estimating larger- compared to smaller-component fractions. This was not the case when they estimated more common small-component fractions relative to other simplified fractions. Adults were more familiar with, and confident in, their estimates of common fractions relative to other simplified fractions, even though they were equally precise in their estimates of both types of fractions. Similarly, in Experiment 2, when adults estimated a different set of fractions and reported their strategy use after each trial, they were equally precise when estimating larger- and smaller-component fractions, even though they were less confident in their estimates of larger-component fractions. One possibility for adults' higher confidence despite being equally precise, is that they might be incorrectly relying on fraction familiarity or other non-diagnostic cues to judge their confidence, rather than their objective performance.

These findings are consistent with an unpublished study from our lab involving children's fraction reasoning. Fifth grade students were more confident (on a scale ranging from 1 to 3) when

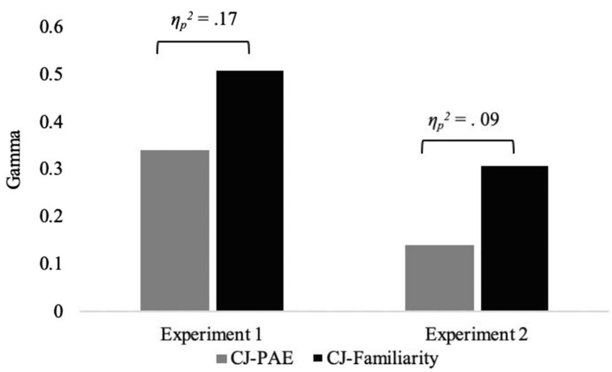

Fig. 1 Differences in gamma for the CJ-PAE relation compared to the CJ-Familiarity relation in Experiments 1 and 2

estimating smaller-component fractions ($M = 2.58$, $SE = .06$) relative to equivalent larger-component fractions ($M = 2.42$, $SE = .07$; $p < .001$, $\eta_p^2 = .36$), even though they were *less* precise when estimating smaller- ($M = 11.3\%$, $SE = 1.5\%$) compared to larger-component fractions ($M = 9.2\%$, $SE = 1.2\%$; $p = .04$, $\eta_p^2 = .08$). Thus, fifth grade students' pattern of confidence in estimating fraction magnitudes was similar to some first grade students' pattern of confidence when estimating whole number magnitudes on a smaller vs. larger numerical scale (Wall et al. 2016). The results from these prior experiments suggest that confidence judgments may be driven by factors other than objective estimation precision, such as familiarity with numbers. The results of the current experiments support this familiarity hypothesis: intra-individual variability in confidence was more strongly related to familiarity ratings than to PAE, even though adults rated their confidence, but not familiarity, immediately after making each estimate. These results have implications for the applicability of cue-utilization theory to metacognitive monitoring in mathematics.

Cues for judging confidence

Individuals use a variety of information to make metacognitive judgments (Koriat 1997). In our work, adults' confidence judgments were more strongly related to their ratings of fraction familiarity than to their objective number line estimation precision. This was the case in *both* experiments, despite differences in the experimental design of the parent study (Fitzsimmons et al. in press). Thus, adults are more confident when they estimate fractions that they rate as more familiar, suggesting familiarity might serve as a cue for judging confidence.

The use of familiarity as a cue for judging confidence aligns with past work on other types of metacognitive judgments, such as feelings-of-knowing or judgments of learning (Alter and Oppenheimer 2009; Hertzog et al. 2003; Koriat 1997, 2008; Koriat and Levy-Sadot 2001; Reder and Ritter 1992). For instance, adults were more likely to say they could recall the answer to an arithmetic problem (e.g., 25×39) instead of indicating they needed to compute the answer, if they had previously been presented problems with similar components (e.g., $25 + 39$; Reder and Ritter 1992). Thus, increasing exposure to the problem components increased adults' feelings-of-knowing without affecting their ability to correctly recall answers to arithmetic problems.

Some research suggests adults use familiarity when making confidence judgments in other domains. Adults' confidence in their ability to make predictions about basketball teams' likelihood of winning or losing a game was influenced by non-diagnostic information about the teams. Adults were more confident in their predictions and less likely to use provided statistical information about wins and losses when team names were displayed, yet they were no more accurate in their predictions, especially if they were fans of the teams displayed (Hall et al. 2007). Thus, additional non-diagnostic cues, such as familiarity with team names, hindered confidence judgment accuracy. In line with past work, adults in both of our experiments provided higher confidence judgments when estimating fractions they rated as more familiar, even when their estimates were not more precise than estimates of less familiar fractions.

Task order and other influences on metacognition

In both of our experiments, all three aspects of fraction reasoning–confidence, familiarity, and PAE–were related (see Table 2). However, participants rated their familiarity at the end of the experiments after completing a number of other tasks. Thus, it is not possible to determine whether they rated their familiarity by reflecting on all previous formal classroom and informal

experiences with these fractions (which likely varied from participant-to-participant), the extent to which they remembered seeing the fractions within this specific study (which was the same for each participant), or both.

Participants' ratings of familiarity might have been influenced by their confidence ratings provided on each trial in the number line estimation task at the beginning of each experiment. However, given prior research on the role of familiarity in confidence and feeling-of-knowing judgments (e.g., Koriat and Levy-Sadot 2001; Reder and Ritter 1992), it seems more likely that had participants rated their familiarity with fractions at the beginning of the experiment, this would have influenced later judgments of confidence. Thus, we made a conscious decision to evaluate participants' familiarity with fractions at the end, rather than at the beginning of each experiment, because we wanted to eliminate the possibility that judging familiarity might influence judging confidence later in the study. By having participants rate confidence immediately after each number line estimation trial before making any other subjective judgments (i.e., familiarity), we aimed to *increase* the strength of the relation between confidence and estimation precision (PAE). It is also possible that adults interpreted our familiarity question as a broad measure of their confidence in working with fractions. This interpretation of our familiarity question seems unlikely because adults in Experiment 1 rated fractions that appeared more frequently in math textbooks (Braithwaite and Siegler 2018) as more familiar than other simplified fractions, suggesting that they interpreted our measure as the extent to which they recognized each fraction on the list. While we have provided an important first step in investigating the relation between familiarity and confidence judgments, future work should examine how task order might influence these relations. Furthermore, we acknowledge that our findings are correlational in nature and cannot confirm the direction of the relation between confidence and familiarity. As such, other factors might have influenced adults' ratings of familiarity and confidence.

There are many factors that can influence adults' metacognitive judgments. For example, fluency or speed of processing (e.g., Ackerman and Koriat 2011; Koriat et al. 20144) as well as beliefs (Mueller and Dunlosky 2017; Mueller et al. 2014; Finn and Tauber 2015) influence monitoring in memory tasks and might influence metacognitive monitoring in mathematics as well. In the current experiments, we did not measure speed of estimation for two reasons. First, response time was controlled in one condition in Experiment 1 for half of the participants. Second, response times vary as a function of the fraction magnitude. It takes longer to correctly place larger fractions on the line because these magnitudes are located further to the right side of the line. In this online study, it would be impossible to know whether longer response times were due to difficulty processing the fraction or due to dragging the slider further across the number line. Additionally, participants implement various strategies to estimate numerical magnitudes, and some strategies take more time to implement than others (Fazio et al. 2016; Sidney et al. 2018). In addition to fluency, beliefs about math might influence confidence judgments.

Adults' beliefs about math involving larger numbers might have influenced their confidence judgments in the current experiments. Adults may have judged their confidence based on their beliefs that math with bigger numbers is typically more difficult, even if it was not always the case in this particular study. In other words, their actual confidence in their estimates of fractions with smaller or larger components might have been equivalent, but they rated their confidence lower when estimating larger-component fractions because they noticed the surface-level differences between component sizes and not the structural similarities in magnitude (cf. Gentner 1983). Thus, they rated their confidence differently across the fraction types. However, even when comparing confidence among more and less common small-component fractions in Experiment 1, adults were more confident and familiar with more

common fractions despite having equal performance. There was no manipulation of component size for these fractions (i.e., they all had single digit numerators and denominators), thus participants could not have responded solely based on the size of components.

To further examine the information people use to judge their confidence, researchers can ask children and adults to provide verbal reports of the reasons *why* they judge their confidence and familiarity in the way that they do. Self-report measures are one successful way to assess the cues people use when making metacomprehension judgments (e.g., Jaeger and Wiley 2014; Thiede et al. 2010). These measures can elucidate whether processes involved in participant responses might contribute to their metacognitive judgments.

Implications, the role of environmental input, and future directions

These findings have implications for numerical reasoning in everyday life. Adults in our sample tended to be both confident and precise when estimating fractions they rated as familiar. The role of familiarity in confidence may be especially pronounced for adults who are less precise when estimating magnitudes. Adults with less precise representations of rational numbers might be more likely to make errors with high confidence when they are reasoning about financial or medical decisions (e.g., Peters et al. 2019). For example, adults who estimate the time it takes to pay off a loan or who estimate their absolute risk of experiencing negative side effects from medical treatments might not check their work if their familiarity with the numbers involved in these calculations influences their confidence in their solutions.

The role of familiarity in confidence may also have implications for children who are just beginning to formally learn about fractions. When learning about fractions, children may be more confident when solving problems involving more familiar fractions, even if they are no more accurate on these problems. Alternatively, children who are immersed in learning about fractions during formal instruction may treat all fractions as equally unfamiliar and thus use other cues to guide their confidence (i.e., amount of time it took them to make their estimate, difficulty of problem). Then, they may begin to use familiarity as a cue for their confidence after additional years of formal instruction. Future work should examine how instruction and experience influence what information, such as familiarity cues or processing fluency, individuals use to judge their confidence in the domain of mathematics. Indeed, analyses of three math textbooks suggests some fractions are more frequently presented than others, and this exposure influences children's reasoning about fraction arithmetic (Braithwaite and Siegler 2018). Preliminary evidence (Eason and Ramani 2018) suggests that parent-child number talk about fractions in formal and informal contexts can be experimentally manipulated, yet it is an open question as to whether environmental input has downstream consequences for individuals' metacognition related to fraction learning and errors.

Research involving whole numbers suggests that environmental input does have consequences for later knowledge: parents with higher socioeconomic status talked more frequently about numbers in their homes than parents with lower socioeconomic status, and number talk predicted cardinal number knowledge (i.e., knowing how many objects are in a set) after controlling for socioeconomic status (Gunderson and Levine 2012; Levine et al. 2010). Further, manipulating environmental input has downstream consequences for whole number knowledge: when children who were enrolled in Head Start played a numeric board game for an hour, their number knowledge improved (Ramani and Siegler 2008; Ramani et al. 2012; Siegler and Ramani 2008). Thus, number knowledge is influenced by environmental input, and environmental input can be manipulated (e.g., more talk, board games, etc.), yet the impact of environmental

experiences on confidence and familiarity with numbers, specifically fractions, remains an open question. Our results suggest that children's experience with fractions commonly presented in textbooks may have downstream consequences for adults' fraction metacognition, but future work should examine whether familiarity and confidence can be experimentally manipulated.

Concluding remarks

When adults are presented with highly familiar math content, they will be more confident in their performance. This higher confidence may, in turn, reduce the likelihood that (1) they will allot additional study time to master familiar, yet difficult course content, or (2) they will check for errors when familiar numbers are used in important decision-making scenarios in everyday life (e.g., financial or health decisions). This research is a first step towards extending theoretical accounts of metacognition to the domain of numerical cognition, suggesting that familiarity with fractions serves as a cue for confidence when reasoning about magnitudes.

Funding information Support for this research was provided in part by the U.S. Department of Education, Institute of Education Sciences grant R305A160295 to Dr. Clarissa A. Thompson and the Kent State University Judie Fall Lasser Award to Charles J. Fitzsimmons.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Appendix

Table 5 Smaller and larger component stimuli presented in Experiment 1 (top) and 2 (bottom)

Experiment 1			
Smaller component		Larger component	
1/5	3/15	5/25	6/30
2/9	4/18	6/27	8/36
1/4	3/12	8/32	9/36
1/3	3/9	10/30	12/36
3/7	9/21	12/28	15/35
1/2	2/4	12/24	15/30
5/9	10/18	15/27	20/36
2/3	4/6	6/9	8/12
3/4	6/8	9/12	12/16
4/5	8/10	12/15	20/25
5/6	10/12	15/18	20/24
Experiment 2			
Smaller component		Larger component	
1/6		12/72	
2/7		24/84	
4/9		36/81	
3/5		27/45	
5/8		35/56	
6/7		54/63	

References

Ackerman, R., & Koriat, A. (2011). Response latency as a predictor of the accuracy of children's reports. *Journal of Experimental Psychology: Applied, 17*(4), 406–417.

Alibali, M. W., & Sidney, P. G. (2015). Variability in the natural number bias: Who, when, how, and why. *Learning and Instruction, 37*, 56–61.

Alter, A. L., & Oppenheimer, D. M. (2009). Uniting the tribes of fluency to form a metacognitive nation. *Personality and Social Psychology Review, 13*(3), 219–235.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01.

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology, 64*, 417–444.

Braithwaite, D. W., & Siegler, R. S. (2017). Developmental changes in the whole number bias. *Developmental Science, 21*(2).

Braithwaite, D. W., & Siegler, R. S. (2018). Children learn spurious associations in their math textbooks: Examples from fraction arithmetic. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 44*(11), 1765–1777.

Braithwaite, D. W., Pyke, A. A., & Siegler, R. S. (2017). A computational model of fraction arithmetic. *Psychological Review, 124*(5), 603–625.

Braithwaite, D. W., Leib, E. R., Siegler, R. S., & McMullen, J. (2019). Individual differences in fraction arithmetic learning. *Cognitive Psychology, 112*, 81–98.

Dehaene, S., & Mehler, J. (1992). Cross-linguistic regularities in the frequency of number words. *Cognition, 43*(1), 1–29.

Dumas, J. E., Johnson, M., & Lynch, A. M. (2002). Likableness, familiarity, and frequency of 844 person-descriptive words. *Personality and Individual Differences, 32*(3), 523–531.

Dunlosky, J., & Metcalfe, J. (2008). *Metacognition*. Sage Publications.

Dunlosky, J., & Rawson, K. A. (2012). Overconfidence produces underachievement: Inaccurate self evaluations undermine students' learning and retention. *Learning and Instruction, 22*(4), 271–280.

Eason, S. H., & Ramani, G. B. (2018). Parent–child math talk about fractions during formal learning and guided play activities. *Child Development., 91*, 546–562. https://doi.org/10.1111/cdev.13199.

Fazio, L. K., DeWolf, M., & Siegler, R. S. (2016). Strategy use and strategy choice in fraction magnitude comparison. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*(1), 1.

Finn, B., & Tauber, S. K. (2015). When confidence is not a signal of knowing: How students' experiences and beliefs about processing fluency can lead to miscalibrated confidence. *Educational Psychology Review, 27*(4), 567–586.

Fitzsimmons, C. J., Thompson, C. A., & Sidney, P. G. (in press). Do adults treat equivalent fractions equally? Adults' strategies and errors during fraction reasoning. Journal of Experimental Psychology: Learning, Memory, and Cognition.

Fleming, S. M., & Lau, H. C. (2014). How to measure metacognition. *Frontiers in Human Neuroscience, 8*, 443.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science, 7*(2), 155–170.

Gunderson, E. A., & Levine, S. C. (2012). Some types of parent number talk count more than others: Relations between parents' input and children's cardinal-number knowledge. *Developmental Science, 14*(5), 1021–1032.

Hall, C. C., Ariss, L., & Todorov, A. (2007). The illusion of knowledge: When more information reduces accuracy and increases confidence. *Organizational Behavior and Human Decision Processes, 103*(2), 277–290.

Handel, M. J. (2016). What do people do at work? *Journal for Labour Market Research, 49*(2), 177–197.

Hertzog, C., Dunlosky, J., Robinson, A. E., & Kidder, D. P. (2003). Encoding fluency is a cue used for judgments about learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(1), 22.

Higham, P. A., & Higham, D. P. (2019). New improved gamma: Enhancing the accuracy of Goodman–Kruskal's gamma using ROC curves. *Behavior Research Methods, 51*(1), 108–125.

Jaeger, A. J., & Wiley, J. (2014). Do illustrations help or harm metacomprehension accuracy? *Learning and Instruction, 34*, 58–73.

Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General, 126*(4), 349–370.

Koriat, A. (2008). When confidence in a choice is independent of which choice is made. *Psychonomic Bulletin & Review, 15*(5), 997–1001.
Koriat, A., & Levy-Sadot, R. (2001). The combined contributions of the cue-familiarity and accessibility heuristics to feelings of knowing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(1), 34.
Koriat, A., Ackerman, R., Adiv, S., Lockl, K., & Schneider, W. (2014). The effects of goal-driven and data-driven regulation on metacognitive monitoring during learning: A developmental perspective. *Journal of Experimental Psychology: General, 143*(1), 386–403.
Levine, S. C., Suriyakham, L. W., Rowe, M. L., Huttenlocher, J., & Gunderson, E. A. (2010). What counts in the development of young children's number knowledge? *Developmental Psychology, 46*(5), 1309–1319.
Metcalfe, J. (2009). Metacognitive judgments and control of study. *Current Directions in Psychological Science, 18*(3), 159–163.
Mueller, M. L., & Dunlosky, J. (2017). How beliefs can impact judgments of learning: Evaluating analytic processing theory with beliefs about fluency. *Journal of Memory and Language, 93*, 245–258.
Mueller, M. L., Dunlosky, J., Tauber, S. K., & Rhodes, M. G. (2014). The font-size effect on judgments of learning: Does it exemplify fluency effects or reflect people's beliefs about memory? *Journal of Memory and Language, 70*, 1–12.
Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin, 95*(1), 109–133.
Nelson, L. J., & Fyfe, E. R. (2019). Metacognitive monitoring and help-seeking decisions on mathematical equivalence problems. *Metacognition and Learning, 14*, 1–21. https://doi.org/10.1007/s11409-019-09203-w.
Ni, Y., & Zhou, Y. D. (2005). Teaching and learning fraction and rational numbers: The origins and implications of whole number bias. *Educational Psychologist, 40*(1), 27–52.
O'Leary, A. P., & Sloutsky, V. M. (2016). Carving metacognition at its joints: Protracted development of component processes. *Child Development, 88*(3), 1015–1032.
O'Leary, A. P., & Sloutsky, V. M. (2018). Components of metacognition can function independently across development. *Developmental Psychology, 55*(2), 315.
Opfer, J. E., & DeVries, J. M. (2008). Representational change and magnitude estimation: Why young children can make more accurate salary comparisons than adults. *Cognition, 108*(3), 843–849.
Peters, E., Västfjäll, D., Slovic, P., Mertz, C. K., Mazzocco, K., & Dickert, S. (2006). Numeracy and decision making. *Psychological Science, 17*(5), 407–413.
Peters, E., Hibbard, J., Slovic, P., & Dieckmann, N. (2007). Numeracy skill and the communication, comprehension, and use of risk-benefit information. *Health Affairs, 26*(3), 741–748.
Peters, E., Tompkins, M. K., Knoll, M. A., Ardoin, S. P., Shoots-Reinhard, B., & Meara, A. S. (2019). Despite high objective numeracy, lower numeric confidence relates to worse financial and medical outcomes. *Proceedings of the National Academy of Sciences, 116*(39), 19386–19391.
R Core Team. (2015). R: A language and environment for statistical computing. In *R Foundation for statistical computing*. Vienna: Austria. URL http://www.R-project.org/.
Ramani, G. B., & Siegler, R. S. (2008). Promoting broad and stable improvements in low-income children's numerical knowledge through playing number board games. *Child Development, 79*(2), 375–394.
Ramani, G. B., Siegler, R. S., & Hitti, A. (2012). Taking it to the classroom: Number board games as a small group learning activity. *Journal of Educational Psychology, 104*(3), 661–672.
Reder, L. M., & Ritter, F. E. (1992). What determines initial feeling of knowing? Familiarity with question terms, not with the answer. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(3), 435.
Sidney, P. G., Thalluri, R., Buerke, M. L., & Thompson, C. A. (2018). Who uses more strategies? Linking mathematics anxiety to adults' strategy variability and performance on fraction magnitude tasks. *Thinking & Reasoning*, 1–38.
Siegler, R. S., & Booth, J. L. (2004). Development of numerical estimation in young children. *Child Development, 75*(2), 428–444.
Siegler, R. S., & Opfer, J. E. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological Science, 14*(3), 237–250.
Siegler, R. S., & Pyke, A. A. (2013). Developmental and individual differences in understanding of fractions. *Developmental Psychology, 49*(10), 1994–2004.
Siegler, R. S., & Ramani, G. B. (2008). Playing linear numerical board games promotes low-income children's numerical development. *Developmental Science, 11*(5), 655–661.
Siegler, R. S., Thompson, C. A., & Schneider, M. (2011). An integrated theory of whole number and fractions development. *Cognitive Psychology, 62*(4), 273–296.
Siegler, R. S., Duncan, G. J., Davis-Kean, P. E., Duckworth, K., Claessens, A., Engel, M., Susperreguy, M. I., & Chen, M. (2012). Early predictors of high school mathematics achievement. *Psychological Science, 23*(7), 691–697.

Tekin, E., & Roediger, H. L. (2017). The range of confidence scales does not affect the relationship between confidence and accuracy in recognition memory. *Cognitive Research: Principles and Implications, 2*(1), 49.

Tekin, E., Lin, W., & Roediger, H. L. (2018). The relationship between confidence and accuracy with verbal and verbal+ numeric confidence scales. *Cognitive Research: Principles and Implications, 3*(1), 41.

Thiede, K. W., Griffin, T. D., Wiley, J., & Anderson, M. C. (2010). Poor metacomprehension accuracy as a result of inappropriate cue use. *Discourse Processes, 47*(4), 331–362.

Torbeyns, J., Schneider, M., Xin, Z., & Siegler, R. S. (2015). Bridging the gap: Fraction understanding is central to mathematics achievement in students from three different continents. *Learning and Instruction, 37*, 5–13.

Wall, J., Thompson, C., & Morris, B. J. (2015). Confidence judgments and eye fixations reveal adults' fractions knowledge. In D. C. Noelle, R. Dale, A. S. Warlaumont, J. Yoshimi, T. Matlock, C. D. Jennings, & P. P. Maglio (Eds.), *Proceedings of the 37th annual meeting of the cognitive science society* (pp. 2571–2576). Austin: Cognitive Science Society.

Wall, J. L., Thompson, C. A., Dunlosky, J., & Merriman, W. E. (2016). Children can accurately monitor and control their number-line estimation performance. *Developmental Psychology, 52*(10), 1493–1502.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.